

Vizsgálat két csoportban

Kérdés: A két minta származhat-e azonos populációból, vagy a két populáció paraméterei azonosak?

paraméteres

$$\mu_1 = \mu_2 ?$$

Nullhipotézis: $\mu_1 = \mu_2$

kétmintás t-próba

nem paraméteres

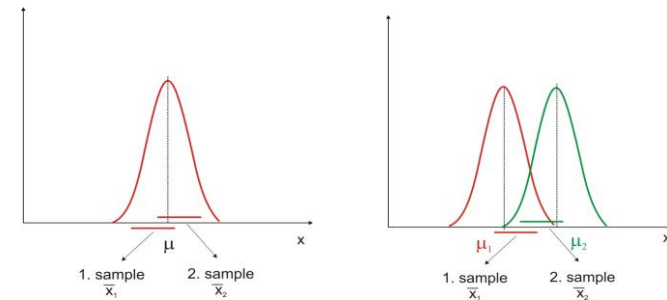
Nullhipotézis: a két minta azonos populációból származik.

Mann-Whitney U-próba

Kétmintás t-próba

egy populáció
(a két átlag eltérése véletlen)

két populáció
(a két átlag eltérése nem véletlen)



Standard hiba

$$s_1 = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n_1 - 1}}$$

$$s_{\bar{x},1} = \frac{s_1}{\sqrt{n_1}}$$

$$s_2 = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n_2 - 1}}$$

$$s_{\bar{x},2} = \frac{s_2}{\sqrt{n_2}}$$

Közös standard hiba: a két standard hiba súlyozott átlaga.

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{Q_1 + Q_2}{n_1 + n_2 - 2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Kétmintás t-próba

$$\bar{x}_1 \neq \bar{x}_2$$



Ismert eloszlású változóra van szükség!



$$s^* = \sqrt{\frac{Q_1 + Q_2}{n_1 + n_2 - 2}}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

A próba

A t-érték az t-érték!

Akkor meg tudom csinálni!
Pardon, mennyi a szabadsági fokok száma?

$$\text{sz.f.} = n_1 + n_2 - 2 \\ ((n_1 - 1) + (n_2 - 1))$$



A kétmintás t-próba feltétele

- A feladat: két egymástól **független** csoport összehasonlítása.
- A változó **normális eloszlású** legyen.
- A **szórás** a két csoportban **azonosnak** tekinthető.

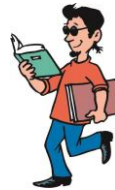


Ez utóbbi új!
Hogyan állapítható meg?

A szórások vizsgálata

Hogyan fogjunk hozzá?

Nullhipotézis: a két szórás azonos, az eltérés véletlen (mintavétel).



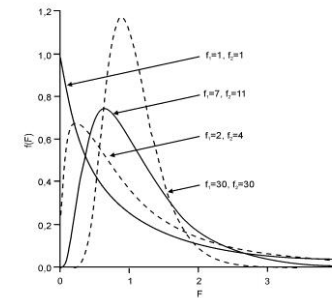
De hiszen ez olyan, mint egy hipotézis-vizsgálat!

Az F-próba

A nullhipotézishez tartozik egy ún. F-eloszlás.

$$F = \frac{s_1^2}{s_2^2}$$

Szabadsági fokok:
számláló: $n_1 - 1$
nevező: $n_2 - 1$



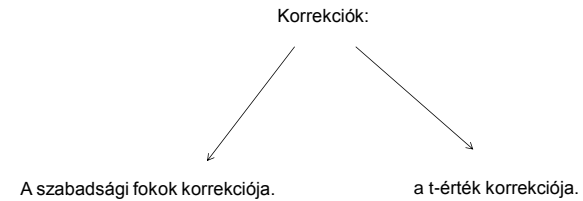
A döntés

- 1. Ha a véletlen eltérés valószínűsége kicsi ($p \leq \alpha$) – **elvetjük** a nullhipotézist.
- 2. Ha a véletlen eltérés valószínűsége nagy ($p > \alpha$) – **megtartjuk** a nullhipotézist.

F táblázat

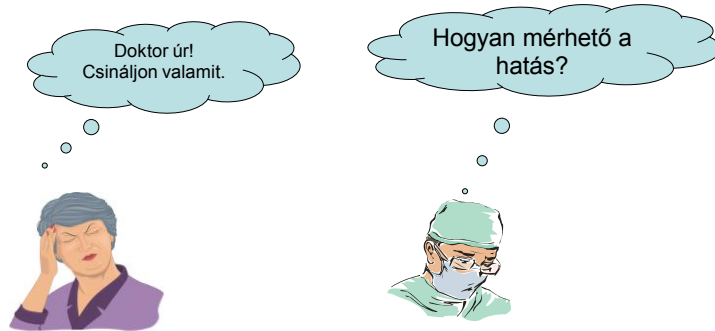
Számítógép: F-próba

Ha a két szórás nem azonos!

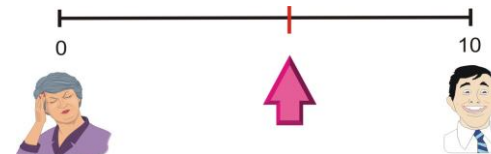


Mann-Whitney U-próba

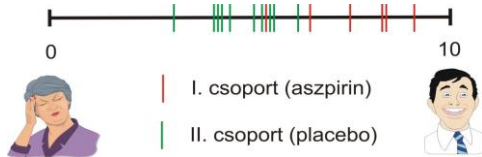
Példa: hatásos-e a fejfájás-csillapító?



Kísérlet

I. csoport:
(eset)
aszpirint kapII. csoport:
(kontroll)
placebo-t kap
(hatóanyag nélküli
tabletta)Ez egy önkényes,
folytonos skála.

Eredmények



érték	3,1	4,1	4,2	4,3	4,5	5,1	5,3	5,4	5,5
rang	1	2	3	4	5	6	7	8	9
érték	5,6	6,2	6,2	6,5	7,5	8,3	8,3	8,4	9,1
rang	10	11,5	11,5	13	14	15,5	15,5	17	18

A nullhipotézis megfogalmazása

a „gyógyszer” nem hatásos.

A két csoport azonos populációhoz tartozik.



A rangok összege (avagy a kis Gauss esete a tanárral)

Gyerekek! Adjátok össze a számokat 1-től százig.

Miért adjam össze? Könnyebben is kiszámolható!



$$1 + 100 = 101$$

$$2 + 99 = 101 \dots$$

$$\sum_{i=1}^n i = \frac{n}{2} \cdot (n+1)$$

A rangok összege

T – a rangok összege az I. csoportban, véletlen eloszlás esetén a várható értéke:

$$n_1 \cdot \frac{n_1 + n_2 + 1}{2}$$

(n_1 elem, amelyek átlaga = $(n_1 + n_2 + 1)/2$)

Nullhipotézis: az ettől való eltérés véletlen.

Kis n : egy U-eloszlás írja le a véletlen eltérés valószínűségét.

A „nagy átalakítás”

Ha n elég nagy:

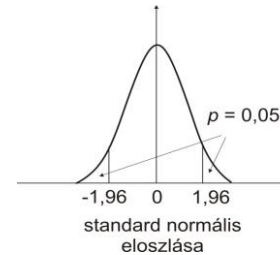
$$z = \frac{T - n_1(n_1 + n_2 + 1)/2}{s}$$

$$s = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}$$

A z változó standard normális eloszlású.



Döntés



A kiszámolt z -érték: 3,24.
Ez nagyobb, mint az 1,96.

Következtetés: a nullhipotézist elvetjük.

Kiszámolt p -érték $< 0,1\%$.
Következtetés hasonló a fentihez.

Variancia-analízis (ANOVA)

A



B



C



Van különbség a csoportok között?

Nincs, az eltérés csak véletlen! Ez a nullhipotézis.



A variancia összetevői

	1. csoport	2. csoport	3. csoport
1	173	170	175
2	175	163	174
3	169	165	171
4	168		172
5			172
csoportátlag	171,25	166	172,8

nagyátlag = 170, 58

$$(170 - 170,58) = (170 - 166) + (166 - 170,58)$$

$$(175 - 170,58) = (175 - 172,8) + (172,8 - 170,58)$$

$$(x_{i,j} - \bar{x}) = (x_{i,j} - \bar{x}_j) + (\bar{x}_j - \bar{x})$$

csoportok közötti különbség

csoporton belüli (pl. véletlen) eltérés

\bar{x} - nagyátlag

\bar{x}_j - csoportátlag

A nullhipotézis

A csoportok között nincs különbség.

A csoportok közötti eltérés csupán a véletlen műve.



Döntés: a csoportok közötti és a csoporton belüli variációk összehasonlítása alapján.



Hogyan hasonlítjuk össze?

Variációk összehasonlítása? Ilyenről már volt szó!

Valóban, a kétmintás t-próba esetében.

$$F = \frac{MS_A}{MS_E}$$



A döntés

- 1. Ha a véletlen eltérés valószínűsége kicsi ($p(F \geq F_{\text{krit}}) \leq \alpha$) – **elvetjük** a nullhipotézist.
- 2. Ha a véletlen eltérés valószínűsége nagy ($p(F \geq F_{\text{krit}}) > \alpha$) – **megtartjuk** a nullhipotézist.

(A döntés után, ha szükségesnek tartjuk, csinálhatunk t-próbákat)

Kruskal-Wallis próba

Ha a változó nem normális eloszlású!

Az adatokat a csoportoktól függetlenül rangsoroljuk!



Rangsorolás

	1. csoport	2. csoport	3. csoport
1	173	170	175
2	175	163	174
3	169	165	171
4	168		172
5			172

elem	163	165	168	169	170	171	172	172	173	174	175	175
rang	1	2	3	4	5	6	7,5	7,5	9	10	11,5	11,5

csoport	elemszám	rangok összege
1	4	27,5
2	3	8
3	5	42,5

Milyen eloszlást használjunk?

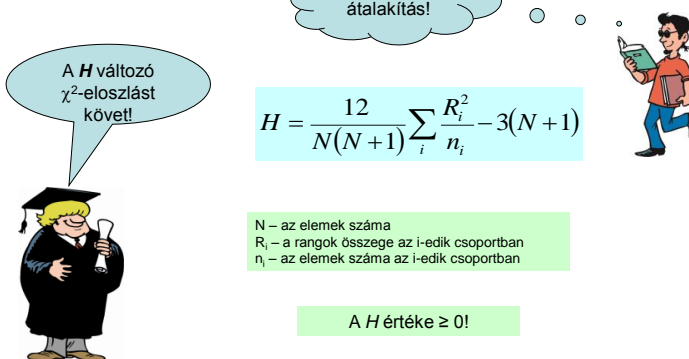
A H változó χ^2 -eloszlást követ!

Akkor jön az átalakítás!

$$H = \frac{12}{N(N+1)} \sum_i \frac{R_i^2}{n_i} - 3(N+1)$$

N – az elemek száma
 R_i – a rangok összege az i -edik csoportban
 n_i – az elemek száma az i -edik csoportban

A H értéke ≥ 0 !



Hasonlítsuk össze!

	1. csoport	2. csoport	3. csoport
1	173	170	175
2	175	163	174
3	169	165	171
4	168		172
5			172

ANOVA

 $\alpha = 0,05$
 $p = 0,024$

Kruskal-Wallis próba

 $\alpha = 0,05$
 $p = 0,083$
Döntés:
 elvetjük a nullhipotézist.

Döntés:
 megtartjuk a nullhipotézist.

 Hipotézis
 vizsgálat?

- Felállítjuk a **nullhipotézist**.
- Keresünk egy **ismert eloszlású változót**.
- Az eloszlás alapján kiszámoljuk a **véletlen eltérés valószínűségét**.
- Ha ez kisebb mint a szignifikancia szint **elvetjük**, ellenkező esetben **megtartjuk a nullhipotézist**.
- Ennyi!

